고정 선반입과 Drop 정책을 이용한 VOD 버퍼 관리 알고리즘 설계 및 평가

박규석[†] · 문병철[††]

요 약

VBR로 압축된 멀티미디어 데이터는 비트 변화율이 매우 큰 편차로 변하기 때문에 자원 예약관리가 매우 힘들다. 따라서 VOD 서버는 MPEG데이터의 참조 패턴을 오프라인으로 분석한 메타 테이블을 근거로 과부하 구간에서 미리 선반입하여 시스템 활용율을 높이는 선반입 기법을 사용한다. 이러한 선반입 기법은 과부하 구간에서도 상영실패는 발생하지 않으나, 선반입 임계슬롯이 증가함으로써 버퍼 점유량이 증가하여 버퍼의 최소 적재시간과 적재비용을 유지할 수 없다. 따라서 본 논문에서는 선반입 구간을 고정적으로 사용하여, 과부하 구간에서도 적재비용과 적재시간을 일정한 범위 이하로 유지하면서 시스템 자원의 활용율을 높이는 방법을 제안한다. 그리고 선반입 구간을 고정시킴으로써 발생하는 상영 실패는 Drop모듈을 사용하여 GOP내의 B프레임에 한정하며, 미디어의 질 저하를 전체 사용자에게 고르게 분산시키는 방법을 제안한다.

Design and Evaluation of a VOD Buffer Management Algorithm Using Fixed Prefetch and Drop Strategics

Kyoo-Seok Park[†] and Beong-Cheoul Moon[††]

ABSTRACT

Resource managing is very difficult because the multimedia data compressed by a VBR shows a bit rate change within high range. Therefore, the VOD server should use the prefetch method in order to improve system utilization, such as prefetching data in the overload period by a meta table that is the reference pattern of MPEG data which is analyzed off line. This prefetch method will not result in a failure to display at overload, however, this method can not keep a minimum loading time and low costs, because the prefetched section is being maximizes.

In this paper, we suggest another method that the system utilization can be improved using the fixed prefetched section to keep loading time and costs under a constant range at overload. But this technique will result in a failure to display, due to fixed prefetched section. But, in this paper we suggest a drop module that drops only the B frame in the GOP, consistently distributes a lower drop in media quality for the user.

1. 서 론

컴퓨터 기술의 비약적인 발전은 인간 문화의 새로운 의사전달로써 텍스트, 이미지, 비디오, 오디오등 모든 지식 전달매체를 통합한 멀티미디어를 탄생하게 하였으며, 처리기의 성능향상, 고속 대용량 장치의 개발 및 컴퓨터 통신망의 발전은 VOD 서버(Video On Demand)와 같은 멀티미디어 응용들이 실생활에서 사용 가능하게 하였다.

이러한 멀티미디어 응용의 일반적인 형태는 멀티미디어 데이터를 저장하고 있는 저장서버와 저장서버에게 원하는 데이터를 요구하는 사용자 시스템 및 이들을 연결하는 통신망으로 구성되는데, 저장 서버가 시스템 성능에 큰 영향을 끼친다. 따라서 멀티미

[†] 종신회원, 경남대학교 공과대학 정보통신공학부 교수
[††] 경남대학교 사회교육원 정보기술교육부 전임연구원

디어 데이터를 저장하고 있는 저장 서버에서의 효율적인 데이터 관리를 위한 많은 연구들이 진행되고 있다.

멀티미디어 데이터는 연속적이면서 동시에 표현되어야 하며, 데이터 량이 매우 크기 때문에 이를 저장, 처리, 전송함에 있어 압축기법 사용은 필수적으로 요구된다. 또한 정해진 시간 내에 처리되지 못할 경우 서비스 품질이 떨어지게 되므로 데이터 전달 경로 상에 필요한 자원들을 예약함으로써 연속매체의 실시간 처리를 보장할 수 있다. 그러나 MPEG과 같은 가변 비트율 압축기법으로 압축된 데이터들은 데이터 내용의 변화로 인해 단위 시간에 재생해야 할 프레임의 크기가 일정하지 않고 이러한 데이터 량의 가변성은 VOD 서버에서 자원 예약관리를 매우 어렵게 한다.

따라서 본 논문에서는 VOD 서버의 효율적인 자원관리를 위하여 가변적으로 변하는 MPEG 데이터의 참조 패턴을 오프라인으로 분석하여 선반입 기법에서 사용함으로써, 가변 데이터로 인한 자원할당의 어려움을 해결한다. 그러나 선반입 임계구간을 산출하는데 있어서 기존의 선반입 기법을 사용할 경우, 과부하 구간에서 선반입 임계슬롯의 증가로 인하여 버퍼 점유량이 증가하여 버퍼의 최소 적재시간과 적재비용을 유지하기 어렵다. 따라서 본 논문에서는 선반입 임계구간을 고정시킴으로써 과부하가 심한 구간에서도 선반입 기법에서 중요한 평가요소인 적재비용과 적재시간을 일정한 범위 이하로 유지하면서 시스템 자원 활용율을 높이는 방법을 제안한다. 그리고 고정 선반입 기법의 사용은 상영실패가 발생할 수 있는 단점이 있으므로 본 논문에서는 상영실패를 MPEG의 복원에 영향이 적은 B프레임에 대한 Drop을 수행하며, Drop모듈을 사용하여 상영실패로 인한 서비스 품질의 저하를 각 사용자에게 분산시키도록 하였다.

2. 관련연구

2.1 MPEG의 구성과 알고리즘

MPEG 비디오 압축 알고리즘은 압축률을 높이기 위해 프레임 내 코딩과 프레임 간 코딩을 사용하고 있다. 프레임 내 코딩만으로 압축한 I프레임은 임의 접근과 고속 찾기와 같은 기능을 지원하는데, 다른 프

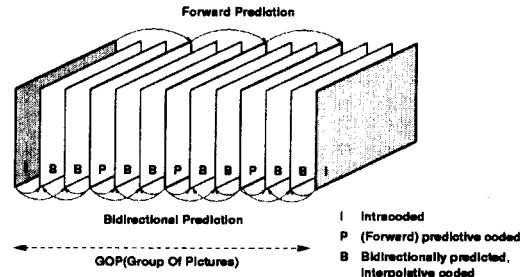



그림 1. MPEG의 구성

레임과 관계없이 독립적으로 압축되므로 단독 복원이 가능한 반면 다른 프레임에 대한 참조 없이 부호화하기 때문에 데이터의 크기가 크다. 아울러 프레임 내 코딩만으로는 압축률을 높일 수 없고, 임의 접근을 쉽게 하기 위하여 무한정 I프레임을 허용할 수는 없다. 이에 대한 절충으로 프레임 내 코딩과 프레임 간의 코딩과의 적절한 균형이 필요하다.

프레임 간 코딩으로 압축된 프레임은 예측 방법에 따라 P프레임과 B프레임으로 나눈다. P프레임은 가장 최근의 I프레임과 P프레임을 참조하여 이전 프레임과 현재 프레임의 차이 값만을 부호화하여 저장하기 때문에 I프레임 보다 크기가 작다. B프레임은 과거 프레임과 미래 프레임을 참조하여 양방향 예측으로 압축하기 때문에 P프레임 보다 압축률이 좋다. 또한 임의접근이나 수정을 위해 독립적으로 복원이 가능한 프레임들의 집합을 MPEG에서는 GOP(Group Of Picture)라 하며, I프레임에서 다음 I프레임의 이전까지의 프레임들로 구성된다. 이러한 MPEG기법으로 압축된 데이터는 데이터 내용의 변화에 따라 가변적인 데이터 크기의 프레임들로 구성된다[10].

2.2 버퍼 관리 기법

VOD 서버에서 버퍼를 관리하는 기법은 크게 각 사용자들에게 독립적인 버퍼를 할당하여 사용하도록 하는 지역 버퍼 관리 기법과 모든 사용자들이 버퍼를 공유하는 전역 버퍼 관리 기법으로 나눌 수 있다. 지역 버퍼 관리 기법의 문제점은 각 사용자마다 버퍼를 할당하고 지역적으로 관리함으로써, 같은 데이터 블록을 참조하는 다른 사용자가 있을 경우에도 이들 사용자간의 데이터를 공유할 수 없다. 그러나 전역 버퍼 관리 기법의 경우 버퍼 공유가 가능하지만, 데이터 소비율이 차이가 나는 사용자들이 함께

버퍼를 사용할 때 데이터 소비율이 높은 사용자의 데이터 블록들이 VOD 서버의 버퍼를 차지할 수 있다. 그러나 여러 사용자들이 데이터 소비율이 비슷하고 데이터 블록에 대한 미래 참조 패턴을 비교적 정확하게 예측할 수 있기 때문에 한 사용자의 데이터 블록에 의해 버퍼가 독점되는 경우는 매우 드물다. 이러한 이유로 최근의 연구들에서는 전역 버퍼 관리 기법을 사용한다[11].

2.3 선반입 기법

선반입 기법은 오프라인으로 분석된 MPEG 데이터의 참조 패턴을 이용하여 디스크의 대역폭에 의한 승인 사용자 제한을 서버의 버퍼에 의한 제한으로 전환함으로써 더 많은 사용자를 수용할 수 있다. VOD 서버가 현재 n명의 사용자를 수용하고 있고, 각 사용자 C_i $(0 \le i \le n-1)$는 서로 다른 영화 M_2를 검색하고 있다고 가정한다. C_2를 위해 한 개의 슬롯 t에서 제공되어야 할 데이터 블록의 수를 $C_2(t)$라고 할 때, 한 슬롯 t에서 모든 사용자들에 의해 소비되는 데이터 블록의 수 $C(t)$는 식 (1)과 같이 나타낼 수 있다.

$$C(t) = \sum_{i=0}^{n-1} C_i(t) \quad (1)$$

한 슬롯동안 VOD 서버가 읽어들일 수 있는 최소의 데이터 블록의 수를 D_{min}, 모든 사용자를 위한 서비스 시작 슬롯이 α, 서비스가 끝나는 슬롯을 ω라 하면 모든 사용자들의 요구를 지연없이 실시간에 처리하기 위해서는 식 (2)와 같은 디스크 대역폭에 대한 조건이 만족되어야 한다.

$$C(t) \le D_{min} \quad (\alpha \le t \le \omega) (2)$$

여기에서 슬롯 $t(\alpha \le t \le \omega)$가 식 (3)을 만족하면 슬롯 t를 영역 $[\alpha, \omega]$에서의 과부하 슬롯이라 한다.

$$C(t) \ge D_{min} \quad (3)$$

이러한 과부하 슬롯에서 데이터 요구량을 이전 슬롯에서 여분의 디스크 대역폭을 이용하여 미리 VOD 서버의 버퍼로 선반입 할 수 있는데 이때 선반입 하는 데이터를 VOD 서버의 버퍼에 적재하는 시간을 최소로 하면서 언제, 얼마나 선반입 해야 할 것인가를 결정해야 한다. 선반입 데이터 블록의 버퍼 적재 시간을 최소화하면서 과부하 슬롯의 데이터 소비량을 만족하기 위해 선반입을 작업을 시작해야 하는 슬롯은 슬롯 $t(\alpha \le t \le \omega)$가 영역 $[\alpha, \omega]$에서의 과부하 슬롯일 때 $(t-\psi+1) \times D_{min} \ge \sum_{i=\psi}^{t} C(t)$를 만족하는 최대의 슬롯 $\psi(\le t)$를 과부하 슬롯 t의 선반입 시작 슬롯이라 한다[1].

2.4 Drop정책

ACME(Abstraction for Continuous MEdia)는 캘리포니아 대학에서 개발한 오디오와 비디오 입출력을 제어하는 네트워크 서버로써, 연속매체의 입출력 및 여러 매체 재생기(스피커, 마이크로 폰, 카메라, 비디오 등)를 제어할 수 있다. ACME에서는 시간 제약을 어기는 경우, 전송될 프레임의 일부를 스키핑(Skipping) 혹은 퍼징(Pausing)하는 기법을 사용하고 있다[3].

HeiTS(Heidelberg Transport System)는 IBM 유럽 네트워킹 센터에서 개발한 디지털 오디오와 비디오의 실시간 전송을 위한 멀티미디어 시스템으로 통신망의 부하에 따른 전송률 제어를 기반으로 하는 시스템이다. 미디어 스케일링(Media Scaling)은 데이터 스트림 본래의 내용을 부표본화(Subsampling)하여 그 데이터 스트림의 일부분만을 나타내는 것을 의미하는 것으로, HeiTS는 수신측 자원의 상태를 동적으로 적용시키기 위하여 수신측 시스템 자원이 과부하 상태에 도달했다면 서비스를 중단하는 것이 아니라 미디어 스트림의 질을 점진적으로 감소시켜 조절함으로써, 사용자에게 의미있는 미디어 스트림을 제공해 준다[7].

QoS 협상에 의한 생략정책은 정적 협상 방식과 서비스 중간에 파라메터를 바꿀 수 있는 동적 협상 방식이 있다. 동적 협상 방식에는 QoS 관리자가 주기적으로 시스템 부하를 검사하여, 필요할 경우 서비스 관리자에게 서비스 품질 협상을 요구하는 운영체제에 의한 방식과 서비스 관리자가 시스템 부하 정도를 질의하여 서비스 품질을 조정하는 방식이 있다[4,5].

저장장치에서 생략정책은 비디오 소스를 MPEG 부호기를 이용하여 프레임을 생성한 후, 프레임의 중요성을 따져 중요도가 높은 I프레임을 추출하여 저장하고, 그 다음 P프레임을 추출하여 저장하며, 마지막으로 B프레임을 추출하여 순차적으로 저장하는 방

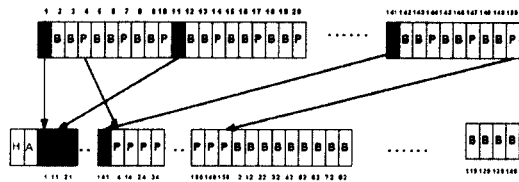



그림 2. MPEG 프레임의 재구성

법을 이용하여 중요도가 높은 순서대로 디스크에 저장한다.

중요도가 낮은 프레임을 Drop시키는 방법으로 142, 143, 145, 146, 148, 149프레임 대신에 I프레임 이전에 위치한 109, 119, 129, 139, 149의 프레임을 생략 전송하여 프레임을 Drop시키는 방법을 제안하고 있다. 이 정책은 클라이언트의 매끄러운 재생을 보장하기 위하여 제안되었고, 참조 프레임이 I와 P프레임이 Drop되는 것을 방지하여 프레임 분실을 개선하는 방안으로 제시되었다[8].

2.5 승인정책

승인 제어 알고리즘은 시스템에 완전한 성능을 보장해주는 결정적인 기법과 절대적인 보장은 해 주지 않지만 가능성이 있는 형식의 관측에 기반한 기법과 양극단 방법의 절충적인 형태인 통계적인 기법이 있다. 결정적인 승인제어 기법은 자원의 이용률이 낮고, 지나치게 승인 제어를 하는 반면 통계적인 승인 제어 기법은 상대적으로 자원의 이용률을 높일 수 있지만 서비스에 대한 보장을 완벽하게 해줄 수는 없다.

연속매체의 상영에서 인간이 감지할 수 없을 정도의 일시적인 성능 저하는 사용자에게 이해되어 질 수가 있을 것이다. 통계적인 승인 제어 기법에서는 새로운 상영이 요청될 때, 그 상영이 평균적으로 요구하는 버퍼 요구량의 적정 α와 기존의 승인된 사용의 예측 버퍼 요구량의 합이 VOD 서버의 버퍼 크기를 넘지 않을 때 요청된 상영은 승인되는데, 여기서 단순한 합과 비교 연산만 있으므로 알고리즘의 자체 복잡도는 $O(1)$이 된다.

그러나 버퍼 요구량의 예측 값이 언제나 VOD 서버의 버퍼 크기를 초과하지 않는다는 보장은 하지 못하므로 일부 스트림들은 특정 슬롯에서 상영되지 못하는 경우가 발생하게 된다. 하지만 앞서 논의한 바와 같이 비디오 상영에서 미세한 량의 상영 실패(Display Fail)는 비디오 전체 서비스에 영향을 거의 미치지 않으므로 사용자가 인내할 수 있는 범위를 벗어나지 않도록 α 값을 조절하면 되고, 이러한 α 값은 결국 시스템의 파라미터이다[2].

3. VOD 서버의 버퍼 스케쥴러

VOD 서버는 수백 편의 영화를 저장하고 있으며 다수의 사용자가 각각 요구한 멀티미디어 데이터를 네트워크를 통하여 멀티미디어 연속성을 유지하면서, 실시간으로 전송해야 하므로 대용량 저장장치와 이를 전송할 수 있는 고속 통신망이 필요하다. 이러한 VOD 서버의 설계는 한정된 자원으로 대용량의 데이터를 처리해야 하는 어려움도 있지만, MPEG으로 압축된 데이터의 비트 변화율이 매우 큰 편차로 변하기 때문에 자원을 할당하는 스케쥴러의 구성을 매우 힘들게 한다. 따라서 본 논문에서는 VOD 서버 자원의 효율적인 운영을 위하여 가변적으로 변하는 MPEG 데이터의 참조 패턴을 오프라인으로 분석한 후, 이를 근거로 과부하가 발생한 영역에서 선반입 기법을 사용함으로써, 가변 데이터로 인한 자원할당의 어려움을 해결한다. 그러나 선반입 임계 구간을 산출하는데 있어 기존의 가변 선반입 기법을 사용할 경우, 과부하 구간에서 선반입과 임계슬롯의 증가로 인하여 최소의 적재시간과 적재비용을 유지하는 것이 어렵다. 따라서 본 논문에서는 선반입 임계구간을 고정시켜 사용함으로써, 과부하 구간에서도 선반입 기법의 중요한 평가요소인 적재비용과 적재시간을 일정한 범위로 유지하면서 VOD 서버의 자원 활용율을 높이는 방법을 제안한다. 그리고 선반입 임계구간의 고정으로 발생할 수 있는 상영 실패를 MPEG 복원에 영향이 적은 B프레임의 Drop에 한정하며, B프레임의 Drop으로 인한 미디어 질 저하를 사용자들에게 분산시키는 Drop 모듈을 설계하였다.

3.1 시스템 구조

본 논문에서 제안하는 고정 선반입 기법은 슬롯 서비스 모듈, 선반입 모듈, Drop 모듈 및 승인제어 모듈을 사용한다. 본 논문에서 제안한 고정 선반입 기법의 시스템 구조는 그림 3과 같다.

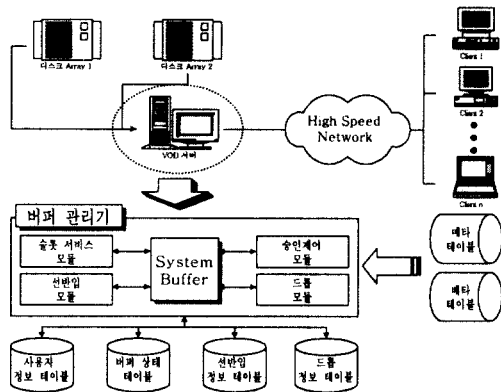



그림 3. 시스템 구조

슬롯 서비스 모듈은 VOD 서버 자원의 효율적인 활용을 위하여 슬롯 단위로 시스템 버퍼를 예약, 할당하고 있으며, 선반입 모듈은 과부하 슬롯이 발생할 경우, 고정된 선반입 구간 내에서 선반입 처리를 수행하고, 승인제어 모듈은 새로운 사용자의 데이터 요구 통계와 서비스 받고 있는 모든 사용자의 데이터 요구의 통계를 분석한다. Drop 모듈은 고정 선반입 구간내에서 선반입이 불가능할 경우, Drop시킬 프레임 선정과 한 사용자에게 과도한 서비스 질 저하를 방지하기 위하여, 상영실패를 각 사용자에게 분산시키는 작업을 수행한다. 이러한 작업을 수행하기 위해서 본 논문에서는 메타 테이블, 사용자 정보 테이블, 버퍼 상태 테이블, 선반입 정보 테이블, 베타 테이블 및 Drop 정보 테이블을 이용한다. 메타 테이블은 VOD 서버에 MPEG으로 저장된 영화의 각 I, B, P프레임의 데이터 크기, GOP 구분 및 Drop 대상 프레임의 우선순위 정보들을 유지하고 있으며, 사용자 정보 테이블은 사용자가 요구한 영화에 대한 정보 및 승인제어를 수행할 때 새로운 사용자의 평균 데이터 요구량 등을 관리하고 있다. 또한 버퍼 상태 테이블은 VOD 서버가 서비스하는 슬롯마다 요구하는 버퍼 요구량을 관리하고 있으며, 선반입 정보 테이블은 버퍼의 용량이이 부족하여 선반입을 수행할 때, 고정 선반입 구간내의 각 슬롯에 대해 선반입을 해야 할 데이터의 정보를 관리하고 있다. 베타 테이블은 각 영화의 각 슬롯에서 영화 종료 슬롯까지의 평균 데이터 요구량을 오프라인으로 분석해 놓은 것으로 승인제어를 수행할 때 새로운 사용자와 기존의 사용자에 대한 데이터 요구량을 분석할 때 사용한다. 그리고 Drop 정보 테이블은 상영실패를 각 사용자에게 분산시키기 위하여 최근에 Drop모듈에 의해 의도적으로 상영 실패된 Drop데이터 소유자의 User_ID와, Drop된 슬롯 및 Drop 량을 관리한다.

3.2 슬롯 서비스 모듈

본 논문에서는 VOD 서버가 여러 사용자에 대한 서비스를 수행하기 위하여 필요한 시스템 자원의 할당을 일정한 시간 단위인 슬롯 단위로 자원을 예약, 할당하는데, MPEG으로 압축된 비디오 데이터의 GOP가 12개의 비디오 프레임으로 구성되었기 때문에 슬롯의 크기는 GOP를 처리하는데 필요한 시간의 정수배가 된다. 따라서 슬롯 서비스 모듈에서는 한 슬롯동안 각 사용자를 위한 자원 예약은 GOP의 배수로 구성할 수 있으며, 이는 시스템 환경에 따라 변경 가능한 시스템 파라메터이다.

한 개의 GOP 프레임 구성은 그림 4와 같다.

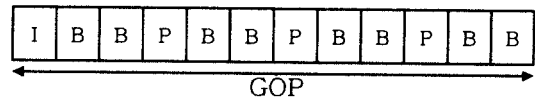



그림 4. GOP의 프레임 구성

VOD 서버가 현재 n명의 사용자를 수용하고 있고, 각 사용자 C_i $(0 \leq i \leq n-1)$는 서로 다른 영화 M_i를 검색하고 있을 경우, 한 사용자 C_i를 위해 슬롯 t동안에 VOD 서버가 버퍼로 읽어 들여야 하는 데이터량 $C_i(t)$는 식 (4)와 같다.

$$C_i(t) = GOP_{M_i} \times k \quad (k=2,4,6\cdots) \tag{4}$$

$C_i(t)$는 다수의 GOP로 구성되는데 k가 클수록 VOD 서버는 데이터를 더 오랫동안 시스템 버퍼에 유지해야 하며, k가 작을수록 승인제어를 위한 알고리즘의 복잡도가 증가하며 VOD 서버가 슬롯 t 동안에 모든 사용자들에 의해 소비되는 데이터 블록의 수 $C(t)$는 식 (1)과 같다.

VOD 서버의 버퍼가 부족할 경우, 선반입을 시작할 슬롯인 선반입 임계슬롯 (ψ)을 구하기 위하여 $(t-\psi+1)\times D_{\min} \geq \sum_{i=\psi} C(t)$를 사용할 경우, 데이터 요구량에 따라 선반입 임계슬롯이 가변적인데, 특히 과부하가 심할 경우 선반입 임계슬롯이 현재 슬롯에서 너무 떨어진 곳에 결정되어 버퍼 점유량이 증가하

여 선반입 기법에서 가장 중요한 요소인 최소의 적재비용과 최소의 적재시간을 유지하기 힘들다. 따라서 본 논문에서는 선반입 임계구간 (ψ)을 상수로 고정하여 시스템 파라메터로 사용하는데, 이는 시스템 환경에 따라 가변적으로 적용 가능하다. 또한 본 논문에서는 고정된 선반입 구간 내에서 요구하는 데이터의 서비스 스케쥴링이 완료한다. 따라서 현재 슬롯이 t 라고 할 때, 현재의 버퍼 상태 $C(t)$는 슬롯 서비스 모듈에서 슬롯 $t-\psi$에서 각 사용자가 요구하는 GOP 내의 프레임의 번호를 사용자 정보 테이블을 통해 읽어 들인 후, 메타 테이블에서 해당 프레임의 데이터 크기를 추출하여 적재, 선반입 혹은 Drop에 대한 서비스 스케쥴링을 구성한다.

3.3 선반입 모듈

슬롯 서비스 모듈에서 $C(t)$를 서비스하면서, 시스템 파라메터인 선반입 임계구간 값 (ψ)을 더한 $C(t+\psi)$의 값을 사용자 정보 테이블, 메타 테이블 및 버퍼 상태 테이블을 이용하여 서비스 스케쥴링을 구성한다. 현재 슬롯 t에서 검사한 슬롯 $t+\psi$의 데이터 요구량 $C(t+\psi)$이 디스크 대역폭을 넘는 경우에는 이러한 슬롯을 과부하 슬롯이라 하며 과부하 슬롯은 식 (5)와 같다.

$$C(t+\psi) \geq D_{min} \tag{5}$$

슬롯 서비스 모듈은 과부하 슬롯이 발생하면 선반입 모듈을 호출하며, 선반입 모듈은 최소의 적재 시간을 유지하기 위하여 $(t+\psi-1)$슬롯에서 $(t+1)$슬롯으로 선반입 가능구간을 검색하게 된다. 이러한 검색은 버퍼 상태 테이블에서 현재 슬롯 (t)에서 상수로 정의된 시스템 파라메터인 선반입 임계슬롯 (ψ)내의 각 슬롯에서 요구한 데이터 총합 ($C(t)$, $C(t+1)$, $\cdots$ $C(t+\psi-1)$)의 정보를 유지하고 있으므로 이를 참조하여 선반입 가능 슬롯을 선택하게 된다.

$$t+1 \leq \text{선반입 가능 구간} \leq t+(\psi-1) \tag{6}$$

본 논문에서 사용하는 버퍼 관리 기법은 2.5절에서 언급한 전역 버퍼 관리 기법을 사용하고 있으며, 식(5)의 디스크 대역폭을 초과하는 데이터 량을 식(6)의 고정 선반입 가능 구간에 선반입 함으로써, 가변 선반입 기법과는 달리 버퍼 요구량을 일정한 범위 이하로 유지할 수 있으며, 선반입 데이터 량으로 인한 스케쥴링의 영향도 가변 선반입 기법에 비해 식(6)의 범위내로 제한시키는 장점이 있다.

그리고 과부하 슬롯이 발생하여 호출된 선반입 모듈은 초과된 데이터의 량 (Δ)는 식(7)에 의해 구해진다.

$$\Delta = C(t+\psi) - D_{min} \tag{7}$$

슬롯 $(t+\psi-1)$에서 선반입의 경우는 두 가지로 나뉘어 지는데, Δ가 슬롯 $(t+\psi-1)$안에 모두 선반입되는 경우 ($\Delta \leq D_{min} - C(t+\psi-1)$)와 Δ의 일부분만 슬롯 $(t+\psi-1)$안에 들어가는 경우 ($\Delta > D_{min} - C(t+\psi-1)$)가 있다. 전자의 경우에는 $t+\psi$에서 발생한 초과 데이터 량이 $t+\psi-1$슬롯에서 모두 선반입함으로써 선반입은 해결되지만, 후자의 경우에는 $t+\psi-2$에서 $t+1$ 슬롯까지 선반입이 계속될 수 있다. 그러나 식(6)의 구간내에 초과된 데이터 량을 모두 선반입하지 못할 경우에는 선반입 모듈은 Drop 모듈을 호출하게 된다.

3.4 Drop모듈

Drop 모듈은 디스크 대역폭을 초과한 데이터 량을 고정 선반입 구간 내에 모두 선반입할 수 없는 경우에 호출되는데, Drop 되는 량을 VOD서버에 서비스 받고 있는 전체 사용자에 분산시킴으로써, Drop발생이 한 사용자에 집중하여 발생하는 것과 Drop으로 인한 급격한 미디어 질 저하를 방지한다.

MPEG에서 I프레임 복원은 자체적으로 복원 가능하며, P프레임은 I프레임을 참조하여 복원되며, B프레임은 I프레임과 P프레임을 참조하여 복원시키는 구조로 되어있다. 따라서 본 논문에서는 미디어 스트림의 질을 보장하기 위하여, I프레임과 P프레임을 Drop 시키지 않으며 GOP내에서 다른 프레임의 복원에 영향을 주지 않는 B프레임만을 Drop시킨다. 특히 B프레임 중 데이터의 크기가 큰 것을 우선적으로 Drop시키는데 이러한 Drop 우선 순위는 오프라인으로 분석하여 메타 테이블에 기록해 놓았다. Drop 프레임의 우선순위는 한 슬롯 동안 처리되는 k개의 GOP중에 B프레임의 크기가 가장 큰 것을 1로 설정하며 그림 5와 같은 자료 구조를 가진다.

그림 5에서 보듯이 B프레임의 크기가 동일한 경우는 한 GOP내에서 연속적인 Drop의 발생을 방지하기 위하여 현재 Drop되는 GOP을 제외한 GOP를 대상

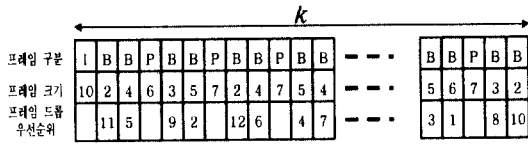



그림 5. Drop 우선순위

으로 왼쪽에서 오른쪽으로 순으로 우선 순위를 설정한다.

Drop 데이터의 소유자 수는 Drop시킬 데이터의 크기 (θ)에 따라 변하며 θ 는 식 (8)에 의해 구해진다.

$$\theta = \Delta - \left(D_{\min} \times \psi - \sum_{k=\psi} \sum_{i=0}^{n-1} C_i(k)\right) \tag{8}$$

VOD 서버에서 Drop모듈이 처음 호출되면 Drop 데이터 소유자의 User_ID는 1번부터 시작되며, Drop 모듈이 한번 이상 호출되었다면 최근에 Drop되었던 데이터의 소유자 User_ID에 1을 더한 User_ID가 선택된다. Drop데이터의 소유자 ($C(i)$)에게 서비스할 k개의 GOP중에서 데이터 크기가 가장 큰 B프레임의 데이터 크기 ($C(i)_{BF_{max}}$)와 θ 의 크기를 비교한다.

$$\theta \le C(i)_{BF_{max}} \tag{9}$$

식 (9)를 만족한다면 사용자 $C(i)$에게 서비스할 GOP의 B프레임 중 데이터 크기가 가장 큰 프레임을 선택하고, 이 데이터는 데이터 적재에 대한 스케쥴링에서 제외된다. 즉 현재 슬롯 t 에서 $t+\psi$ 슬롯에서 선반입 여부를 스케쥴링하여 Drop시켜야 할 데이터가 발생하면 Drop시킬 데이터를 선정한 후, $t+\psi$ 슬롯에서 Drop될 데이터를 제외한 데이터를 적재한다. 아울러 식 (9)을 만족하지 못한다면 식 (10)을 만족할 때 까지 Drop 데이터의 소유자 수를 증가시킨다.

$$\theta \le \sum_{i=0}^{n-1} C(i)_{BF_{max}} \tag{10}$$

3.5 승인제어 모듈

결정적인 승인제어 기법은 서비스 질을 완벽하게 보장해 주지만, VOD 서버의 자원을 효율적으로 사용하지 못하는 단점이 있으며, 관측에 기반한 승인제어 기법은 시스템의 자원을 효율적으로 사용할 수 있지만, 서비스 질을 보장할 수 없다. 한편 통계적인 승인제어 기법은 결정적인 승인제어 기법과 관측에 기반한 승인제어 기법과의 절충적인 효과를 나타내지만, 결정적인 승인제어 기법에 비해 미디어 스트림의 일부에 대한 상영실패가 나타날 수 있다.

그러나 초당 24프레임이 사용자 앞에 나타났다가 사라지는 상황에서, 연속적으로 상영실패가 발생한다면 사용자가 감지할 수 있지만, 일정한 간격을 두고 B프레임 1~2개 정도의 상영실패는 사용자가 감지하지 어려울 뿐만 아니라, 설사 감지된다하더라도 허용될 수 있으리라 본다. 아울러 VOD 서버는 시스템을 효율적으로 사용하며, 시스템 자원의 활용율 및 동시 재생 스트림의 최대화를 위해서도 통계적인 승인제어 기법이 바람직한 알고리즘으로 받아 들여 질 것이다. 한편, 본 논문에서는 주요한 VOD 서버의 성능 평가 항목은 승인 사용자의 수가 아니라 버퍼 점유율이다. 따라서 통계적인 승인 제어기법에서 사용하는 디스크로부터 매체 억세스 시간의 분포와 매체 재생의 요구사항에 대한 평균 및 최악의 값에 대한 분포상태를 고려하지 않고, 메터 테이블을 근거로 한 매체 재생의 데이터 요구량을 단순 산술식을 승인제어를 평가에 사용함으로써 승인제어 복잡도가 $O(1)$이 되도록 하여, 승인제어가 한 슬롯안에 끝날 수 있음을 가정하고 있다[2].

3.5.1 새 사용자의 통계 데이터 요구량 (α) 산출

새로운 사용자 C_n이 요구한 영화 M_i 의 마지막 슬롯이 m이고, 상영을 요청한 시작 슬롯이 t 일 때 α 값은 식 (11)과 같이 구한다.

$$\alpha = \frac{\sum_{i=t}^{m} C_n(i)}{m-t} \tag{11}$$

α 값은 새로운 사용자가 t 슬롯에서 상영을 요청한 영화의 마지막 슬롯인 m까지의 슬롯 당 요구하는 평균 데이터 요구 량으로써 베타 테이블에서는 각 슬롯마다 α 값이 오프라인으로 분석되어 있다. 승인제어 모듈에서는 새로운 사용자가 상영을 요청한 슬롯의 α 값을 베타 테이블에서 참조하여 새로운 사용자에 대한 평균 데이터 요구량으로 사용한다.

3.5.2 승인된 사용자의 통계적인 버퍼 요구량 (β)산출

VOD 서버로부터 서비스 받고 있는 전체 사용자에 대한 버퍼 요구량을 검사하기 위하여 각 사용자의 데이터의 요구량은 베타 테이블을 참조하여 구한다.

베타 테이블에서 참조한 값은 식 (11)에서 언급한 α 값으로 승인된 모든 사용자의 통계적인 버퍼 요구량 (β)는 식 (12)에 의하여 구한다.

$$\beta = \sum_{i=0}^{n-1} \alpha_i \qquad (12)$$

승인제어 모듈은 새로운 사용자가 요구하는 데이터의 평균 값 (α)와 승인제어를 통과하여 VOD서버에 의해 서비스 받고 있는 각 슬롯의 데이터 평균값 (β)의 합이 식 (13)을 만족해야만, 새로운 사용자에 대한 승인제어가 허용된다.

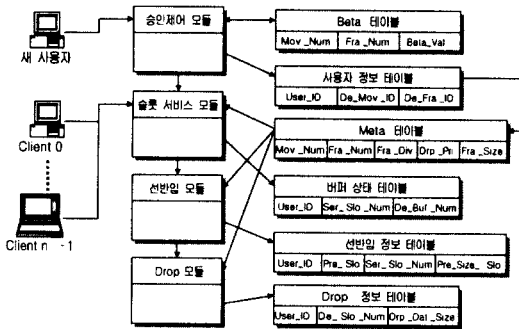
$$\alpha + \beta \le D_{\min} \qquad (13)$$

4. 분석 및 평가

본 논문에서 성능 평가 환경을 구성하기 위하여 오라클 DBMS, 클라이언트 서버용 프로그램 개발 도구인 파워빌더와 Pro*C/C++을 이용하였으며, 고정 선반입 기법의 구성은 그림 6과 같다.

그림 7은 고정 선반입 기법의 동작 화면이다.

그림 6. 고정 선반입 기법의 구성

그림 7. 고정 선반입 기법의 동작 화면

본 논문에서 제안한 고정 선반입 버퍼 관리 알고리즘의 성능 평가 환경은 사용자 요구가 임의로 도착하여 영화를 선택하면 각 기법에 의해 데이터를 적재하는데 이때, 데이터의 적재는 정해진 시간 내에 디스크로부터 VOD서버의 버퍼에 데이터를 적재를 할 수 있음을 가정하였다.

또한 승인제어를 위한 계산은 3.5절에서 언급한 바와 같이 단순 산술식을 사용하므로써 한 슬롯내에서 이루어진다고 가정하였으며, VOD서버와 연결된 통신망은 가변 비트율의 데이터를 전송할 수 있음을 가정하였다. 이는 가변 비트율의 데이터를 통신망으로 전송하기 위해서는 과부하 영역의 시작 시점과 지속기간과 같은 가변 비트율 데이터 흐름의 특성을 통신망에 전달하게 되고, 통신망에서는 이러한 데이터를 기반으로 하여 가능한 데이터의 전송 경로를 설정되므로써 데이터들은 VOD 서버의 버퍼에서 지연됨이 없이 바로 전송할 수 있기 때문에 슬롯 t 에서 소비되는 VOD 서버의 버퍼들은 슬롯 $t+1$에는 데이터의 적재가 가능하다고 가정한 것이다[1].

4.1 가변 선반입 기법과 고정 선반입 기법의 평가

본 논문에서 제안한 고정 선반입 기법의 성능 평가를 위하여 20명의 사용자에게 한 슬롯 당 2 GOP를 서비스하면서 선반입 구간을 10슬롯으로 고정시키고, 디스크 대역폭은 한 슬롯에서 20명의 사용자가 요구하는 평균 데이터 요구량이 약 7.15 MB 로써, MPEG데이터의 데이터 표준편차가 큰 것을 감안하여 8MB 로 설정하였다.

아울러 성능 비교 대상 기법은 다양한 버퍼 관리 기법과 성능 비교를 할 수 있겠으나 본 논문에서 제안하는 고정 선반입 기법과의 비교는 디스크 대역폭 등 시스템 자원을 제한하는 기법과의 비교가 타당하므로 가변 선반입 기법과 성능비교를 하였으며, 가변 선반입 기법의 환경은 디스크 대역폭은 8MB, 20명의 사용자에게 한 슬롯 당 2GOP를 서비스하면서 선반입 임계슬롯을 가변적으로 하였다.

본 논문에서 제안한 고정 선반입 기법과 가변 선반입 기법의 버퍼 점유량은 그림 8과 같다.

고정 선반입 기법은 버퍼 최대 점유량이 11.42MB 까지 증가하는데 비해 가변 선반입 기법은 버퍼 최대 점유량이 33.88MB 까지 증가하는데, 이는 가변 선반입 기법은 과부하 발생 구간이 집중적으로 발생하면

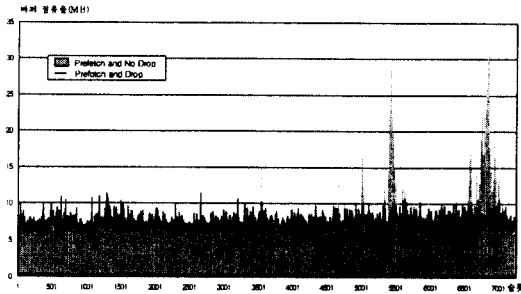



그림 8. 가변 선반입과 고정 선반입 기법과의 버퍼 점유량

선반입 임계구간이 증가되면서 버퍼 점유량이 증가하기 때문이다. 한편 고정 선반입 기법은 과부하가 발생하지 않거나 산발적으로 발생하는 구간에서는 가변 선반입 기법과 동일한 성능을 발휘하며, 과부하 구간이 집중적으로 발생하는 구간에서는 전체 데이터 요구량에 0.31%를 Drop하고, 버퍼 적재 비용을 가변 선반입 기법에 비해 3배 정도로 줄이고 있다. 표 1은 가변 선반입 기법과 고정 선반입 기법에 대한 비교표이다.

표 1. 가변 선반입 기법과 고정 선반입 기법의 항목별 비교

	가변 선반입 기법	고정 선반입 기법
전체 데이터 요구량	51.713 GB	51.713 GB
전체 선반입량	431.8 MB (0.83%)	274.1 MB (0.53%)
전체 Drop량	-	157.7 MB (0.31%)
선반입 최대 구간	275 Slot	9 Slot
선반입 평균 구간	14.8 Slot	3.1 Slot
버퍼 평균 점유량	8.23 MB/Slot	7.29 MB/Slot
디스크 대역폭 평균 활용량	7.13 MB/Slot	7.11 MB/Slot

4.2 선반입 임계슬롯 (ψ)의 평가

표 2는 시스템 파라메터로 설정된 선반입 임계구간의 변화가 시스템 자원 활용율에 미치는 영향을 분석한 것으로 한 슬롯에서 서비스되는 GOP의 갯수는 2개로 설정되었다.

선반입 구간이 증가하면 선반입 량이 증가하고, Drop량은 감소된다. 그러나 선반입 구간이 증가하면 버퍼 평균 점유율이 높아지는데 이는 선반입 기법에서 중요한 요소인 버퍼 적재시간이 증가되었음을 의미한다.

선반입 구간을 10슬롯에서 20슬롯으로 증가시킨 결과 Drop량을 59%감소시키면서, 최대 버퍼 점유율은 16%증가되는데, 선반입 구간의 크기와 버퍼 점유율에 대한 적당한 trade-off가 필요하다.

4.3 Drop 모듈의 평가

각 사용자에게 슬롯 당 2 GOP, 선반입 임계 구간(ψ)을 20슬롯으로 서비스한 결과 사용자의 총 데이터 요구량은 50.5 GB, 총 선반입 량은 338.5 MB (0.65%), 총 Drop량은 93.2 MB(0.18%)이다. 한 사용자의 평균 데이터 요구량은 2586 MB이며, Drop 발생수, Drop량 및 서비스 받은 량은 표 3과 같다.

그림 9는 사용자별 Drop발생의 분포를 나타낸 것

그림 9. 사용자별 Drop발생 분포

표 2. 임계슬롯 ψ의 변화에 대한 성능평가

선반입 임계슬롯	10	20	30
선반입 량	274.1 MB (0.53%)	338.52 MB (0.65%)	368.27 MB (0.71%)
Drop 량	157.7 MB (0.31%)	93.21 MB (0.18%)	63.45 MB (0.12%)
버퍼 평균 점유량	7.29 MB	7.42 MB	7.52 MB
버퍼 최대 점유량	11.42 MB	13.36 MB	16.71 MB
디스크 대역폭 평균 활용율	7.1051 MB/Slot	7.1159 MB/Slot	7.1200 MB/Slot

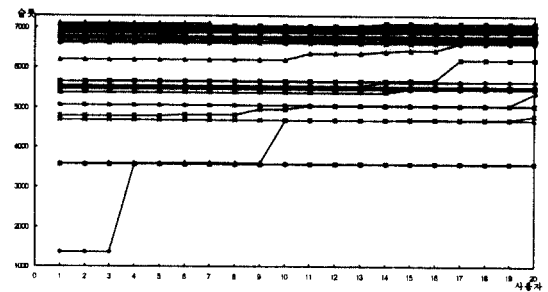



표 3. 사용자별 Drop발생 수, Drop양 및 서비스 받은 양

사 용 자	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Drop 발생 수	49	49	49	49	49	49	49	48	48	48	48	48	48	48	48	48	48	48	48	48
Drop량 (MB)	5.1	6.04	4.16	4.5	4.98	4.54	6.24	3.49	5.1	4.95	4.39	3.7	4.32	5.72	4.53	4.29	3.92	4.98	4.15	4.12
서비스 받은 량(%)	99.8	99.77	99.84	99.83	98.81	99.82	99.76	99.87	99.8	99.81	99.83	99.86	99.83	99.78	99.82	99.83	99.85	99.81	99.84	99.84

이다. 실험결과 Drop 발생간격은 최소 1초에서 최대 36.8분이며, Drop 평균 발생간격은 82.92초마다 B프레임 한 개 이상이 Drop되었다.

5. 결 론

VBR로 압축된 멀티미디어 데이터의 비트 변화율이 매우 큰 편차로 변하기 때문에 VOD 서버에서의 자원관리는 매우 어렵다. 따라서 MPEG데이터의 참조 패턴을 오프라인으로 분석한 후, 과부하 구간 이전에 미리 선반입 하는 기법을 사용하지만 선반입 임계슬롯이 증가하면서 버퍼 점유량이 증가하여 버퍼의 적재 시간과 적재 비용을 일정하게 유지하지 못한다. 따라서 본 논문에서는 선반입 구간을 고정하여 사용함으로써 버퍼 점유량을 감소시키고, 선반입 구간의 고정으로 인하여 발생하는 상영실패를 Drop 모듈을 이용하여 미디어 질 저하를 승인된 전체 사용자에게 분산시킴으로써 한 사용자에 대한 집중적인 Drop으로 인한 미디어 질 저하를 방지하는 기법을 제안하였다. 제안한 고정 선반입 기법은 과부하가 일시적으로 발생하는 구간에서는 가변 선반입 기법과 동일한 성능을 나타내지만, 과부하가 연속적으로 발생하는 상황에서는 전체 서비스해야 할 데이터 량의 약 0.3%을 Drop시키고, 버퍼 요구량을 3배 정도 줄이는 성능을 보여주고 있다. 아울러 Drop된 데이터는 MPEG의 복원에 영향을 주지 않는 B프레임이다. 이러한 B프레임이 각 사용자에게 약 1분마다 Drop되었으며, 이는 미디어 품질의 과도한 저하를 초래하지 않을 것으로 본다.

향후의 연구내용은 고정 선반입 기법을 기반으로 버퍼 공유 기법과 VCR기능을 부가하는 것에 대한 계속적인 연구가 필요하다.

참 고 문 헌

[1] 김순철, "주문형 비디오 서버에서의 효율적인 버퍼 관리 기법" 서울대학교 박사학위 논문

[2] 조경운, "라운드를 바탕으로 한 비디오 서버에서의 버퍼 할당 기법" 서울대학교 석사학위 논문

[3] R. Govindan, David P. Anderson, "Scheduling and IPC Mechanism For Continuous Media", ACM, selected Area in Computer, pp. 68-80, 1991

[4] H. M. Vin, P. Goyal, Anshuman Goyal and Pawan Goyal, "An observation-Based Admission Control Algorithm for Multimedia Server". ICMCS'94, Boston, pp234-243, May, 1994

[5] H. M. Vin, P. Goyal, Anshuman Goyal, "A Statistical Admission Control Algorithm for Multimedia Server", In Proceeding of the ACM Multimedia '94, San Francisco, pp33-40, October, 1994

[6] H. M. Vin, P. V. Rangan, "Admission Control Algorithm for Multimedia On-Demand Servers", The 3rd International Workshop on Network and Operating System support for Digital Audio and Video, Nov, 1992

[7] L. C. Wolf, R. G. Herrtwich, "The System Architecture of the Heidelberg Transport System", ACM Operating System Review, Vol. 28, No. 2, pp.51-64, April, 1994

[8] H. J. Chen, "A Disk Scheduling and MPEG Data Layout Policy Interactive Video Access from a Single Disk Storage Device", Ph. D Thesis, Boston University, 1995

[10] F. Arman, A. Hsu, M. Y. Chiu, "Image Processing on Encoded Video Sequence", Multimedia System, Vol. 1, No. 5, pp 211-219, March, 1994

[11] Banu Ozden, Rajeeve Rastogi, Avi Silberschatz, "Buffer Replacement Algorithms For

Multimedia Storage Systems", In Proceeding of the International Conference on Multimedia Computing and Systems, pp. 172-180, June, 1996

문 병 철

1993년 경남대학교 전자계산학과 졸업(학사)
1995년 경남대학교 대학원 전자계산학과 졸업(공학석사)
1999년 경남대학교 대학원 컴퓨터공학과 박사과정 수료
1996년~현재 경남대학교 사회교육원 정보기술교육부 전임연구원
관심분야 : 멀티미디어 시스템, 시뮬레이션, 분산처리 시스템

박 규 석

1990년~1991년 미 UCLA/객원교수
1992년~1997년 경남대학교 전산정보원 원장
1995년~1996년 한국정보과학회 이사 · 영남지부장
1998년~현재 행정자치부 정책자문위원(전자정보 분과)
1982년~현재 경남대학교 정보통신공학부 교수
관심분야 : 운영체제, 분산처리 시스템, 멀티미디어 시스템, 정보통신 시스템